Cora 2.4

Logiciel d'analyse statistique

Corrélation

Mode d'emploi

Ce logiciel vient compléter la série des logiciels d'analyse statistique édités à mon initiative, qui comprend *Movira, Norma, Alice, Bonnie et Daisy*. Il s'agit cette fois de tester l'hypothèse d'une corrélation linéaire entre deux variables. La notion de corrélation est généralement bien connue du public, mais les conditions dans lesquelles on peut établir ce type de relation le sont moins.

Il est facile d'obtenir sur une calculette scientifique les paramètres d'une droite de régression, même à partir de données non corrélées. Dans ce cas, les nombres obtenus n'ont évidemment aucun sens. C'est pourquoi, avant de s'intéresser à une droite de régression, il faut s'assurer qu'il existe une corrélation entre les deux variables. Deux méthodes sont possibles : le test des signes (test non paramétrique) et le calcul du coefficient de corrélation. La seconde méthode est plus précise, mais elle exige qu'au moins l'une des deux séries soit distribuée normalement.

Une calculette scientifique permet d'obtenir le coefficient de corrélation r. Un étudiant en première année de licence sait que c'est un nombre compris entre -1 et +1, et qu'il n'y a pas de corrélation quand $r = 0$. Mais personne ne lui a dit à partir de quelle valeur de r il y a vraiment une corrélation. Certains choisissent prudemment 0,5 en valeur absolue. C'est souvent une erreur, car tout dépend de la taille de l'échantillon. La limite de r est donnée par une table statistique due au mathématicien Ronald Fisher.

La méthode utilisée par *Cora* consiste à calculer le coefficient de corrélation, puis à le comparer à la table de Fisher. Si r est supérieur en valeur absolue à la limite indiquée, il y a corrélation. Nous avons choisi sur la table un risque d'erreur de 5 % (c'est celui qu'on prend habituellement).

Le mode opératoire est expliqué à l'écran. Il ne faut pas oublier de sélectionner d'abord l'effectif commun aux deux séries. Les données peuvent être des nombres fractionnaires, positifs ou négatifs. La virgule doit être remplacée par un point.

On peut afficher le diagramme de dispersion en allant dans le menu.

Ce test n'est valable que si l'une des deux séries est distribuée normalement. Il est fortement conseillé de s'en assurer par une méthode graphique (voir page suivante). Dans le cas contraire, on peut utiliser le test des signes.

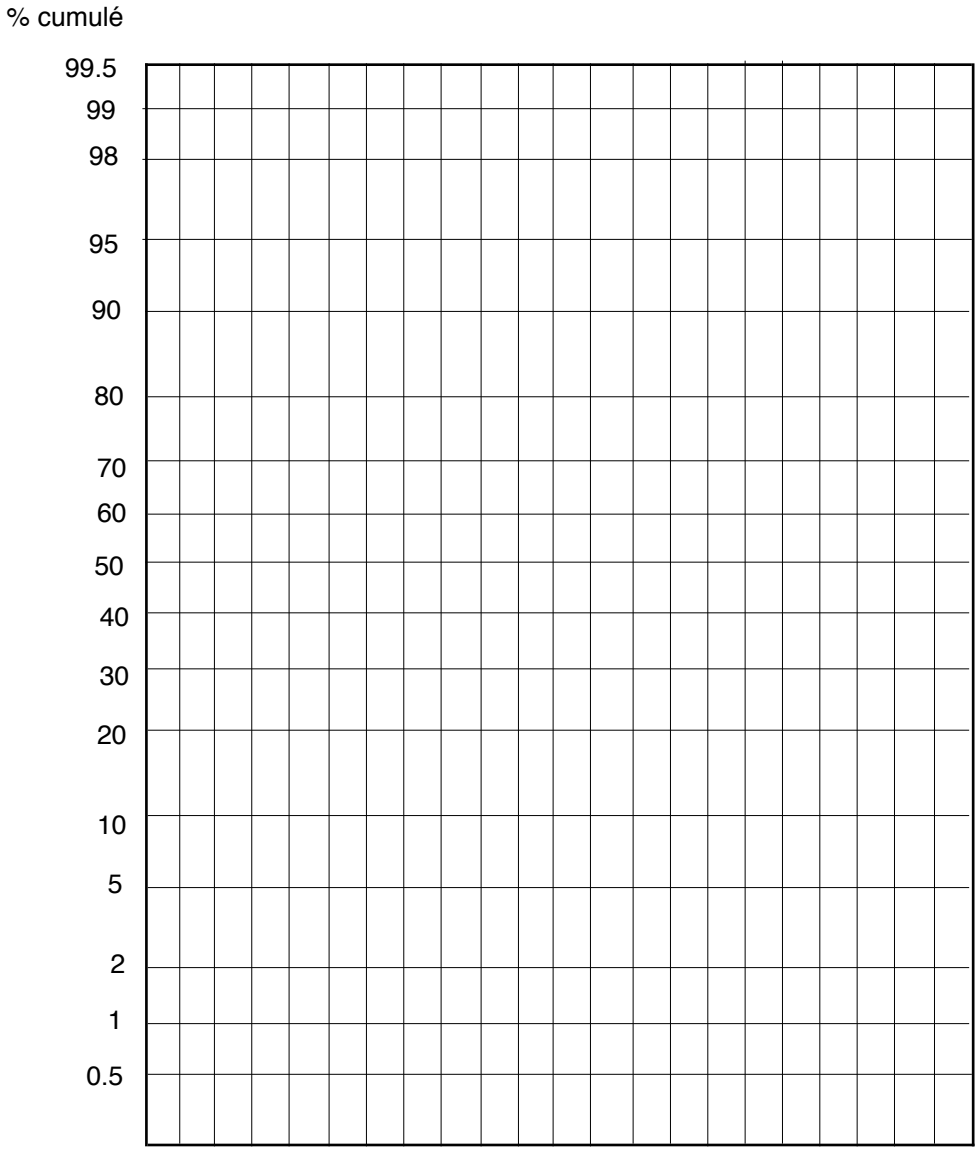
L'effectif de l'échantillon peut être compris entre 5 et 24. L'expérience montre que ces valeurs couvrent la plupart des besoins, du moins dans la vie courante.

Jean-Marie Gogue

Première édition: mai 2004

Nouvelle édition: octobre 2013

Graphique à échelle de Henry



Résultats de mesure (échelle horizontale à déterminer)